Depth Sensation Enhancement for Multiple Virtual View Rendering

Jianjun Lei, *Member, IEEE*, Cuicui Zhang, Yuming Fang, *Member, IEEE*, Zhouye Gu, *Member, IEEE*, Nam Ling, *Fellow, IEEE*, and Chunping Hou

***Abstract*—Depth information is an indispensable element in depth image-based rendering (DIBR) for three-dimensional (3-D) display. In this paper, we propose a novel depth sensation enhancement method to address the problems in multiple virtual view rendering. First, as the depth sensation is decreased when rendering intermediate multiple virtual views, the basic principle of depth sensation enhancement is derived according to the number of rendering views. Second, with the increase of the scene complexity, it is difficult to ensure the depth sensation of all neighboring objects. The saliency analysis is adopted to give preferred guarantee to the depth sensation between the salient object and its neighbors. Then, the depth sensation enhancement for multiple virtual view rendering is performed based on a defined energy function built by the number of rendering views and the saliency analysis. Finally, considering the temporal consistency between adjacent frames, the depth sensation enhancement is extended to video applications with a newly designed energy function with energy term of temporal consistency preservation. Experimental results on a public database demonstrate that the proposed method can obtain promising performance in depth sensation.**

***Index Terms*—Depth image-based rendering (DIBR), depth sensation enhancement, just noticeable depth difference (JNDD), multiple virtual view rendering.**

I. Introduction

THREE-DIMENSIONAL (3-D) display technology has received much attention in recent years due to its improved visual experience to viewers as compared with the traditional 2-D display systems. Many types of 3-D display systems have appeared, including 3-DTVs, handheld game consoles, video conferencing, 3-D cameras, and 3-D mobile phone [1]–[3]. Compared to the traditional 2-D display systems, 3-D display systems provide an additional perception of depth. With depth cameras, depth map can be obtained more easily and accurately. Accordingly, various depth map-based applications emerge and become hot topics in the research community [4], [5].

Virtual view rendering, defined as generating additional views of the scene based on one or several original views, is one of the most important technologies to implement in 3-D displays [6]–[8]. With additional depth information, depth image-based rendering (DIBR) requires much less storage space and transmission bandwidth, and thus becomes a popular technology in virtual view rendering [9], [10].

One common problem in DIBR is that the human eyes cannot perceive the change when the depth difference between objects is less than a threshold. Depth sensation enhancement, which can be used to address the problem, plays an important role in 3-D perception [11]–[13]. Yang *et al.* [14] evaluated the depth sensation ability of observers to distinguish relative depths between objects in a 3-D image presented on a 3-D display. The small size of binocular disparity for background can be used to improve the human eyes' perception ability for relative depth of objects. In [15], a stereoacuity function which describes the just noticeable disparity difference was designed to enhance the depth quality of stereoscopic images. The disparity map is divided into two parts including the coarse disparity layer and the detailed disparity layer. Then, the detailed disparity layer is adaptively manipulated in depth under the guidance of the stereoacuity function. Nur *et al.* [16] investigated the effect of depth map spatial resolution on depth perception and video quality by encoding with different qualities. The depth perception and the video quality are improved when the spatial resolution of the depth map increases. A computational stereo camera system was presented in [17], where the control loop from capture is closed and the physical parameters are automatically adjusted for a better 3-D perception. One of the significant depth sensation enhancement methods is designed based on just noticeable depth difference (JNDD) model. The JNDD indicates that if the depth difference between neighboring objects is smaller than the JNDD threshold, the human eyes cannot perceive the change [18], [19]. Thus, the JNDD based depth sensation enhancement algorithm aims to increase the depth differences between neighboring objects in a visual scene. With depth enhancement operation, human eyes can perceive the depth difference between neighboring objects, which cannot be perceived before enhancement. In [11], a global depth sensation enhancement algorithm was developed

Manuscript received August 08, 2014; revised November 23, 2014 and January 26, 2015; accepted January 29, 2015. Date of publication February 05, 2015; date of current version March 13, 2015. This work was supported in part by the Natural Science Foundation of China under Grant 61271324, Grant 61471262, Grant 91320201, Grant 61471260, and Grant 61202266, and by the Natural Science Foundation of Tianjin under Grant 12JCYBJC10400. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Cha Zhang.

J. Lei, C. Zhang, and C. Hou are with the School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China (e-mail: jjlei@tju.edu.cn).

Y. Fang is with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, China (e-mail: fa0001ng@e.ntu.edu.sg).

Z. Gu and N. Ling are with Santa Clara University, Santa Clara CA 95053 USA (e-mail: guzh0001@gmail.com; nling@scu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMM.2015.2400823

based on the JNDD model. The depth map is segmented into multiple layers by using the kernel density estimation (KDE), and the depth differences between layers are stretched via energy minimization. In [20], an improved JNDD-based depth sensation enhancement algorithm was implemented via an energy minimization framework based on three defined energy terms: depth data preservation, depth order preservation, and depth difference expansion. A modified JNDD measurement method was proposed in [21]. The physical size of objects is adjusted to maintain the perceived size of objects. Then, the JNDD measurement method is applied to depth sensation enhancement, as described in [20]. In those studies, the human depth perception was effectively improved. However, to the best of our knowledge, there is no study investigating the problems of depth sensation enhancement for multiple virtual view rendering. Multiview 3-D display is valuable for providing vivid and immersive 3-D effects. However, lots of 3-D contents just have one or two views of color images plus depth maps. Thus, it is necessary to render multiple virtual view images at the interpolated positions with the available color images plus depth maps. Depth sensation reduction is inevitable in the process of multiple virtual view rendering for 3-D display.

In this paper, we propose a novel depth sensation enhancement method for multiple virtual view rendering. The contributions of the paper mainly include the following. 1) Our work is the first attempt to address the problems in depth sensation enhancement for multiple virtual view rendering, and provides an optimization solution based on the energy function built by the number of rendering views and the saliency analysis; 2) based on the fact that the depth sensation of adjacent objects decreases greatly with the increase of the number of generated intermediate virtual views, we derive the fundamental principle of depth sensation enhancement for multiple virtual view rendering; 3) due to the different complexities of various scenes, we analyze the limitation of depth sensation enhancement according to the number of depth layers, and propose to adopt saliency extraction to preferentially ensure a certain level of depth differences between the salient object and its neighbors in complex scenes; 4) by taking the number of rendering views and the saliency information into consideration, an energy function is designed to implement the depth sensation enhancement for multiple virtual view rendering; and 5) considering the temporal consistency between adjacent frames, we extend the depth sensation enhancement method to video applications with a newly designed energy function with a temporal consistency preservation term.

The rest of the paper is organized as follows. Section II introduces the background and related work. Section III presents the proposed depth sensation enhancement method in detail. We evaluate the performance of the proposed method in Section IV. The final section concludes the paper.

II. Background and Related Work

A. Virtual View Rendering

Virtual view rendering is defined as the process of generating nearby views from original views. The spatial relationship between the original views and generated virtual views based on a rectified multiview setup is shown in Fig. 1. Assume N represent the number of rendering intermediate virtual views, the parallax from the virtual image to left image is inversely proportional to $N+1$. Let $P(X,Y,Z)$ represent a point in the 3-D space. P_L, P_R, and P_i are the projection points of $P(X,Y,Z)$ from different viewpoints. C_L and C_R represent the optical centers of the left camera and the right camera. X_L, X_R, and X_i represent the horizontal coordinates of the projection points in the left viewpoint, the right viewpoint, and the rendered ith virtual viewpoint, respectively. B is the baseline and f is the focal length. The relationship between the virtual views and original views can be described as follows [22]–[24]:

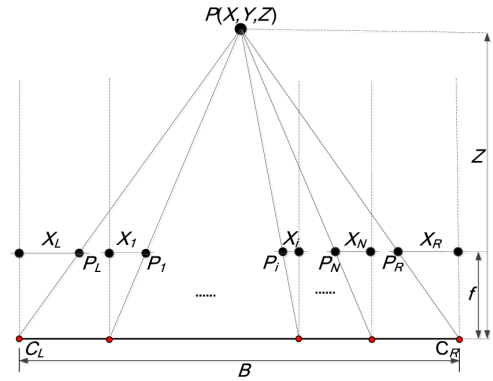


Fig. 1. Spatial relationship between the original views and generated virtual views.

$$p^{Li} = X_L - X_i = c \cdot p^{LR} = \frac{i \cdot p^{LR}}{N+1} = \frac{i \cdot Bf}{(N+1)\,Z} \tag{1}$$

where p^{Li} denotes the disparity between the left and the rendered ith virtual images; p^{LR} is the disparity of the corresponding point pair from the left and right images; c is the scaling factor representing the position of the intermediate virtual viewpoint.

Thus, the intensity value I_i of the pixel in the generated intermediate virtual view can be expressed as follows:

$$I_i\left(X_i, Y\right) = I_i\left(X_L - c \cdot p^{LR}, Y\right) = I_L\left(X_L, Y\right) \tag{2}$$

where I_L is the intensity value of the pixel in the left image.

B. Depth Sensation Enhancement Based on JNDD

Depth sensation enhancement aims to stretch the depth difference between adjacent objects in visual scenes. Jung *et al.* proposed a depth sensation enhancement algorithm by using JNDD [20]. The algorithm firstly divides the depth map into different depth layers by a manual segmentation method. Based on the defined three terms of depth data preservation, depth order preservation and depth difference expansion, an energy minimization framework is designed to stretch the depth difference between adjacent depth layers. The energy function in that study [20] is defined as

$$E\left(m, m^0\right) = E_D\left(m, m^0\right) + E_O\left(m, m^0\right) + E_{JNDD}\left(m, m^0\right) \tag{3}$$

where m^0 and m denote the sets of average depth of the original and enhanced depth objects, respectively; E_D is used to enforce minimal variety of the depth values in the depth image where the direction of depth changes is controlled by an additional weighting function; E_O is used to control the inversion of the local and global depth orders between objects. The most important JNDD term, E_{JNDD}, is adopted to increase the depth differences between objects, and is defined as follows:

$$E_{JMDD}\left(m, m^0\right) = w_1 \sum_{i=1}^{M} \sum_{j \in \Psi_i} \max\left(0, D_{JMDD}\left(m_i^0\right) - |m_i - m_j|\right)^2 \quad (4)$$

where w_1 represents a weighting parameter, Ψ_i denotes a set of segment indices which are connected to the ith segment. D_{JNDD} is the minimal difference between objects that can be perceived by the human visual system (HVS), and is given as follows [20]:

$$D_{JNDD}(d) = \begin{cases} 21, & \text{if } 0 \le d < 64, \\ 19, & \text{if } 64 \le d < 128, \\ 18, & \text{if } 128 \le d < 192, \\ 20, & \text{if } 192 \le d \le 255 \end{cases} \quad (5)$$

where d represents the pixel value in the depth map; 255 and 0 are the maximum and minimum depth values in the recorded scene. Generally, the JNDD model is related to the types of display devices, and thus it should be refined with different display environments [18], [20].

The genetic algorithm or the Levenberg-Marquardt (L-M) algorithm can be applied to find the optimal solution in Eq. (3) [11], [20]. The depth value $d^*(i,j)$ of the resultant depth map, which belongs to the kth object, is linearly updated as

$$d^*(i,j) = d^0(i,j) + \left(m_k - m_k^0\right) \quad (6)$$

where (i,j) is the coordinates of the depth value; $d^0(i,j)$ denotes the original depth value at the location (i,j); k represents the number of depth layers, ranging from 1 to M.

With the existing depth sensation enhancement algorithms [11], [20], depth perception is enhanced, and the enhanced depth map can be used for various depth-based multimedia processing applications. However, those methods only stretch the depth difference between adjacent depth layers without considering the number of virtual viewpoints and the importance of different objects in images. Generally, the depth difference of adjacent viewpoints decreases greatly with the increase of the number of virtual viewpoints when intermediate virtual view images are generated. Furthermore, since there are different depth layers in various visual scenes, not all depth differences between neighboring objects in complex scenes can reach JNDD, which limits the enhancement performance.

In this paper, we propose a novel depth sensation enhancement method for multiple virtual view rendering, to address the aforementioned problems. We design a novel depth sensation enhancement operation by considering the number of intermediate virtual viewpoints in visual scenes. Meanwhile, inspired by the widely accepted perceptual theory that human beings' attention is attracted by the saliency information in visual scenes, we propose to use saliency measure to further optimize the depth sensation enhancement operation. Furthermore, the depth enhancement method is extended to video applications considering the temporal consistency. Experimental results demonstrate that the proposed algorithm can obtain better performance than the existing ones.

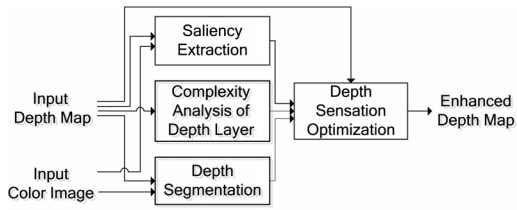


Fig. 2. Framework of the proposed depth sensation enhancement method.

III. Proposed Depth Sensation Enhancement Method

The framework of the proposed method is depicted as Fig. 2. The proposed method mainly includes saliency extraction, complexity analysis of depth layer, depth segmentation, and depth sensation optimization.

Since we aim to control the depth difference of the depth segments to enhance depth sensation, the development of segmentation algorithm is not the goal in this study. Instead, the depth segments are assumed to be available before applying the proposed algorithm. Manual, semiautomatic, or automatic segmentation algorithms [25]–[28] can be used for the depth segmentation. Here, in order to obtain accurate segmentation results, manual segmentation method is mainly adopted.

In the following subsections, we introduce the proposed depth sensation enhancement algorithm in detail.

A. Principle of Depth Sensation Enhancement for Multiple Virtual View Rendering

In order to obtain good visual experiences of depth sensation, it is necessary to enhance the depth difference in 3-D display based on multiple virtual view rendering. In this section, a threshold called MVR_{JNDD} is derived, which relates to the human perceived depth difference between adjacent objects in 3-D display based on multiple virtual view rendering.

As can be observed in Eq. (1), the disparity of each virtual image is reduced greatly when rendering intermediate multiple virtual views. Accordingly, the depth difference between neighboring objects cannot reach JNDD, since the depth difference reduces with the increase of the number of rendering virtual images.

The real-world depth z is usually stored as inverted depth data [23], and the relationship between the value of the normalized depth map d and the real-world depth value can be described as

$$d = round\left(255 \cdot \left(\frac{1}{z} - \frac{1}{z_{\max}}\right) / \left(\frac{1}{z_{\min}} - \frac{1}{z_{\max}}\right)\right) \quad (7)$$

where z_{max} and z_{min} represent the original maximum and minimum depth value of the real-world scene, respectively.

Then, according to Eq. (1), z can be calculated by cBf/p. Similarly, z_{min} and z_{max} are computed by cBf/p_{max} and

cBf/p_{min}, respectively. The relationship between d and p can be expressed as

$$\begin{aligned} d &\cong 255 \cdot \left(\frac{p}{c \cdot B \cdot f} - \frac{p_{\min}}{c \cdot B \cdot f}\right) / \left(\frac{p_{\max}}{c \cdot B \cdot f} - \frac{p_{\min}}{c \cdot B \cdot f}\right) \\ &= \frac{255}{p_{\max} - p_{\min}} \cdot p - \frac{255 \cdot p_{\min}}{p_{\max} - p_{\min}} \\ &= \alpha \cdot p + \beta \end{aligned} \tag{8}$$

where p_{max} and p_{min} represent the maximum and minimum values in the disparity map.

It can be observed from Eq. (8) that β is a constant. As D_{JNDD} represents the just noticeable depth difference between the adjacent objects, the difference of depth can be approximated as α times of the difference of disparity. Taking the parameter $\boldsymbol{\alpha}$ and the number of rendering virtual viewpoints N into consideration, the JNDD model for multiple virtual view rendering, MVR_{JNDD}, can be expressed as follows:

$$MVR_{JNDD}(p) = \frac{(N+1)}{\alpha} \cdot D_{JNDD}(d) \tag{9}$$

According to the MVR_{JNDD} model, those disparity differences between neighboring layers, which lower than MVR_{JNDD}, are required to be stretched.

B. Selective Enhancement for Complex Scenes

The depth sensation enhancement operation tries to enhance the depth difference between each pair of neighboring depth layers to reach JNDD. With the increase of scene complexity, the depth sensation enhancement method cannot ensure the depth difference of each neighboring depth layer. Research studies on visual attention have shown that human eyes do not treat the content equally in a scene and usually focus on some specific regions selectively. Therefore, a limiting factor K is introduced to control the employment of visual attention for depth sensation enhancement in different scenes. The factor K is formulated as

$$K = \frac{256}{(N+1) \cdot \max\limits_{d \in [0,255]} (D_{JNDD}(d))} \tag{10}$$

When the number of depth levels of a scene is less than or equal to K, the scene is defined as a simple scene, and can be directly stretched using the MVR_{JNDD} model. When the number of depth levels is larger than K, the scene is defined as a complex scene, and can be adjusted by the following selective enhancement method. In our experiments, as the same with [20], the depth levels are obtained by depth segmentation. In other words, the number of the depth levels is equal to the number of the layers of depth segmentation.

Considering that the biological vision systems tend to find the most informative region in a scene [29]–[32], visual saliency detection can be adopted to extract salient object in a complex scene. The salient map is computed by using the method in [29], and the most salient segment is marked as the salient region. Then, the depth map is divided into salient and non-salient parts, processed by different strategies. Fig. 3 shows an example for describing the selective enhancement method for the depth map. In this figure, *Obj1* represents the salient object, and

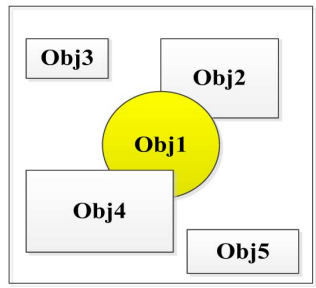


Fig. 3. Example of describing the selective enhancement method.

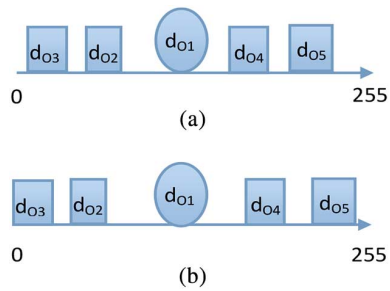


Fig. 4. An instance of the result of the selective enhancement method. (a) Before enhancement. (b) After enhancement.

Obj2 and *Obj4* are the neighboring objects of *Obj1*. *Obj3* and *Obj5* denote other objects which are not neighboring with *Obj1*. If the depth difference between *Obj1* and *Obj2* or *Obj1* and *Obj4* is less than MVR_{JNDD}, we adjust the depth map such that neighboring objects, *Obj1* and *Obj2* or *Obj1* and *Obj4*, in the depth image can have a depth difference which is at least MVR_{JNDD}. Fig. 4 shows a more detailed instance of the result of the proposed selective enhancement method. In this figure, $d_{O1}, d_{O2}, d_{O3}, d_{O4}$, and d_{O5} represent the average depth values of *Obj1*, *Obj2*, *Obj3*, *Obj4*, and *Obj5*, respectively. Assume that the depth differences between *Obj1* and its neighboring objects are less than MVR_{JNDD}, the enhancement result can be seen in Fig. 4(b). The distances, between the salient object *Obj1* and the neighboring objects *Obj2* and *Obj4*, are firstly ensured to be enhanced. The other parts of neighboring objects are also enhanced without ensuring the distance reaching MVR_{JNDD}. Selective enhancement of the depth map is clearly advantageous, since it is impossible to ensure the depth difference of each neighboring object when there are many objects in a scene.

C. Depth Sensation Optimization Based on Energy Function

In this section, we propose an energy function for depth sensation enhancement which takes the multiple virtual view rendering and selective enhancement into consideration. Let x_i^0 represent the average disparity value of the ith depth layer of the original depth map. Similar to (3), the target value for each object is found by minimizing the following energy function:

$$E(x, x^0) = E_D(x, x^0) + E_O(x, x^0) + E_{MVR}(x, x^0) \tag{11}$$

$$x^* = \arg\min_x E(x, x^0) \tag{12}$$

where

$$x^0 = \left\{x_i^0 | i = 1, \ldots M\right\}, x^* = \left\{x_i^* | i = 1, \ldots M\right\} \quad (13)$$

where x^0 and x^* are the sets of the average disparity values of the original and enhanced depth segments, respectively. E_{MVR} enables the stretching of the depth difference, and is defined as follows:

$$E_{MVR}\left(x, x^0\right) = w_1 \sum_{i=1}^{M} \sum_{j \in \Psi_i} S_i \cdot \max\left(0, MVR_{JNDD}\left(x_i^0\right) - |x_i - x_j|\right)^2 \quad (14)$$

where Ψ_i denotes a set of indices of the segments connected to the ith segment; w_1 is a weighting factor. The cost for each depth layer is weighted according to the saliency of the depth layer, and the weight S_i is defined as

$$S_i = \begin{cases} \lambda & \text{if the } i\text{th segmention is salient region,} \\ 1 & \text{otherwise.} \end{cases} \quad (15)$$

where the factor λ is a constant larger than 1. We use a large value of λ for the salient region to ensure the enhancement of the depth differences between salient part and its neighboring objects. According to the limiting factor K, the depth difference between the salient object and the neighboring objects, is given preferred guarantee. The depth differences of other parts are enhanced by optimization algorithms, which cannot ensure the depth differences of all neighboring objects.

The solution of Eq. (12) is generated by using the genetic algorithm, which is a powerful tool for solving nonlinear and nonconvex objective functions.

D. Extension to 3-D Video

In this part, we extend the depth sensation enhancement to video applications with a newly designed energy function. The main difference between video and image is the temporal coherence which exists between continuous frames of a video sequence. Therefore, the temporal coherence constraint must be considered when processing a depth video by depth sensation enhancement algorithm. We propose a novel energy function which takes temporal coherence into account. Differing from the above energy function, a temporal consistency preservation term is added. The target depth for each depth frame is found by minimizing the following energy function:

$$E\left(x(t), x^0(t)\right) = E_D\left(x(t), x^0(t)\right) + E_O\left(x(t), x^0(t)\right) + E_{MVR}\left(x(t), x^0(t)\right) + E_{TMP}\left(x(t), x^0(t)\right) \quad (16)$$

$$x^*(t) = \arg\min_{x(t)} E\left(x(t), x^0(t)\right) \quad (17)$$

where

$$x^0(t) = \left\{x_i^0(t) | i = 1, \ldots M\right\}, x^*(t) = \left\{x_i^*(t) | i = 1, \ldots M\right\} \quad (18)$$

where $x^0(t)$ and $x^*(t)$ are the sets of the average disparity values of the original and enhanced depth segments for the current frame, respectively. E_{TMP} is the temporal consistency preservation term which is defined as follows:

$$E_{TMP}\left(x(t), x^0(t)\right) = \sum_{i=1}^{M} w_i(t) \cdot \left(\left(x_i(t) - x_i^0(t)\right) - \left(x_i^*(t-1) - x_i^0(t-1)\right)\right)^2 \quad (19)$$

where

$$w_i(t) = c_i(t) \cdot \exp\left(-\left|x_i^0(t) - x_i^0(t-1)\right|\right) \quad (20)$$

$c_i(t)$ represents the count of pixels in the ith segment. If the absolute difference between $x^0(t)$ of the current frame and $x^0(t-1)$ of the same position in previous frame is smaller, the confidence of E_{TMP} is higher. Therefore, the error between consecutive frames can be reduced. This temporal term is useful to enforce minimal change of the depth values between adjacent frames.

IV. Experimental Results

A. Experimental Setup and Parameter Setting

In order to evaluate the effectiveness of the proposed method, the color images and the corresponding ground-truth disparity maps provided by Middlebury database [33]–[35] were used in our experiments. Test images include Aloe, Baby1, Venus, Moebius, Midd1, and Bowling1. The focal length is 3,740 pixels and the baseline is 160 mm. In the experiments, the existing holes in the original depth maps are filled using the median filter. The color images are shown in Figs. 5(a)–(f), and the corresponding depth segmentation results, obtained by manual segmentation, are shown in Figs. 5(g)–(l).

The input color image and the consistent depth map were used to render multiple virtual views. The positions of the rendered intermediate virtual views are obtained according to linearly change c in Eq. (2) from 0 to 1 in the experiment. For example, if we render two intermediate views, the values of c are 1/3 and 2/3. The application of the depth sensation enhancement for multiple virtual view rendering is shown in Fig. 6. The enhanced depth map includes the depth map processed by conventional algorithm [20] and the proposed method. All the depth maps used to generate multiview stereoscopic images are preprocessed by an additional depth map filtering method [36] to reduce the holes in the generated virtual viewpoint images.

The solution of Eq. (12) is dependent on the factor λ. The performance of the proposed algorithm is evaluated with respect to the setting of λ when rendering two intermediate views. Fig. 7 shows the number of depth pairs between the salient part and its adjacent objects which depth differences below the MVR_{JNDD} by varying the weighting parameter λ. As λ increases, the influence of salient part in the MVR_{JNDD} term strengthens, and therefore the depth sensation between the salient part and its adjacent objects is preferentially enhanced. When λ is larger than 50, the number of depth pairs between the salient part and its adjacent objects, which depth differences below the MVR_{JNDD}, becomes zero. Considering that a larger λ will influence the enhancement of depth sensation for

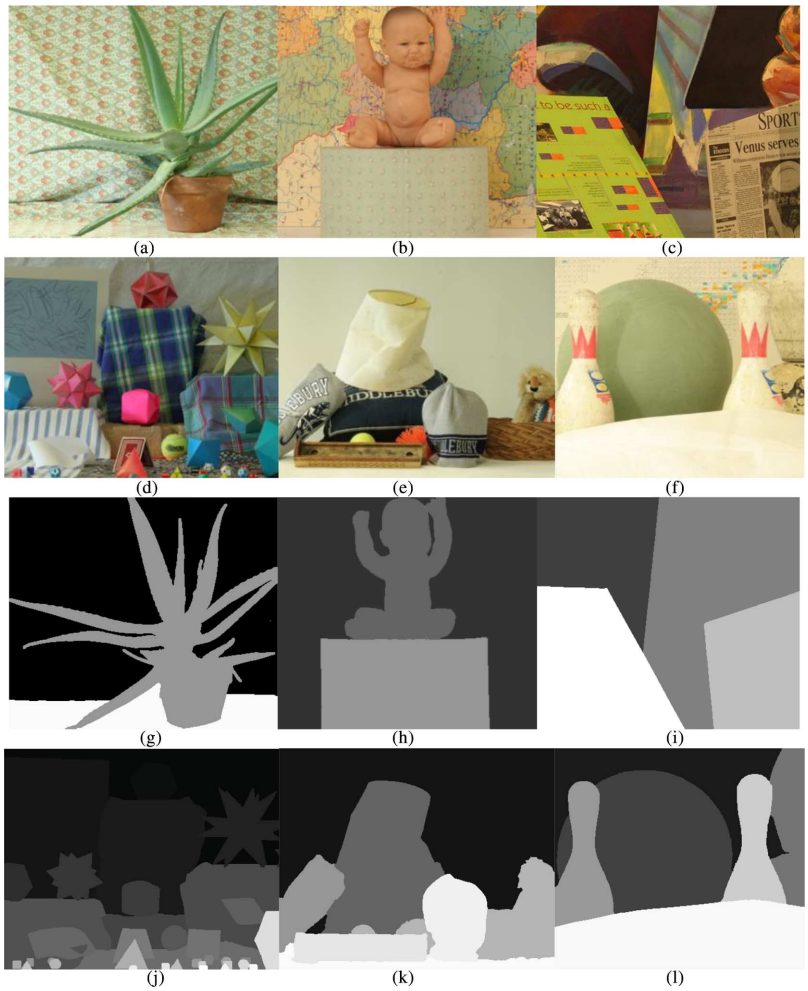


Fig. 5. Monoscopic color images (a)–(f) and the corresponding depth segmentation results (g)–(l). (a) Aloe (427 × 370). (b) Baby1 (413 × 370). (c) Venus (434 × 383). (d) Moebius (463 × 370). (e) Middl1 (465 × 370). (f) Bowling1 (417 × 370).

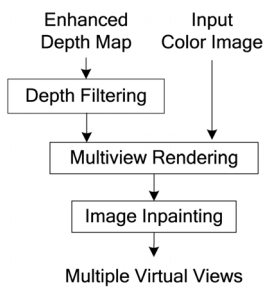


Fig. 6. Application of the depth sensation enhancement for multiple virtual view rendering.

non-salient parts, we empirically set the value of λ to be 50. The other parameters are set as the same with [20].

B. Performance Evaluation on Depth Map and Virtual View

In this section, we first analyze how the energy function modifies the depth map. It should be noted that the proposed method changes the depth difference between depth layers only when the depth difference is less than MVR_{JNDD}. Fig. 8 shows the results of depth sensation enhancement. For comparison, the original depth maps and the resultant depth maps using the conventional method of [20] are also shown in Fig. 8. In this figure, the left column shows the original depth maps, the middle column shows the enhanced results by the conventional method [20], and the right column shows the enhanced results by the proposed method. Obviously, we can observe that the depth differences between many neighboring depth layers which are less than MVR_{JNDD} have been increased. This increase can result in the depth sensation enhancement for the stereoscopic images. Meanwhile, it can be observed that the depth differences between neighboring depth layers are more obvious than those from the conventional method of [20]. We can evaluate the performance of the enhanced depth maps based on the gray-difference. In Aloe, according to the gray-difference between the flower and the background, the right one processed by the proposed method is stronger than the middle one processed by the conventional method. Similarly, the gray-difference between the green paper on the left and the newspaper on the right in Venus, the pink hexagon in the middle and the plaid boxes on the right in Moebius, the gray hat and the yellow board in Middl1, and the green ball and the two bowling pins in Bowling1, all have been increased. It is worthwhile to point out that the pink hexagon in (d), the yellow hat in (e), and the green ball in (f), which are marked with

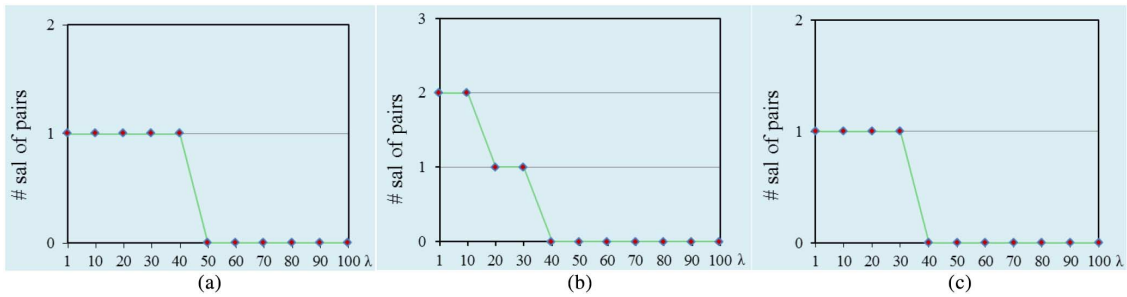


Fig. 7. Number of depth pairs between the salient part and its adjacent objects which depth differences below the MVR_{JNDD} for different values of λ. From left to right are the results for (a) Moebius, (b) Middl1, and (c) Bowling1.

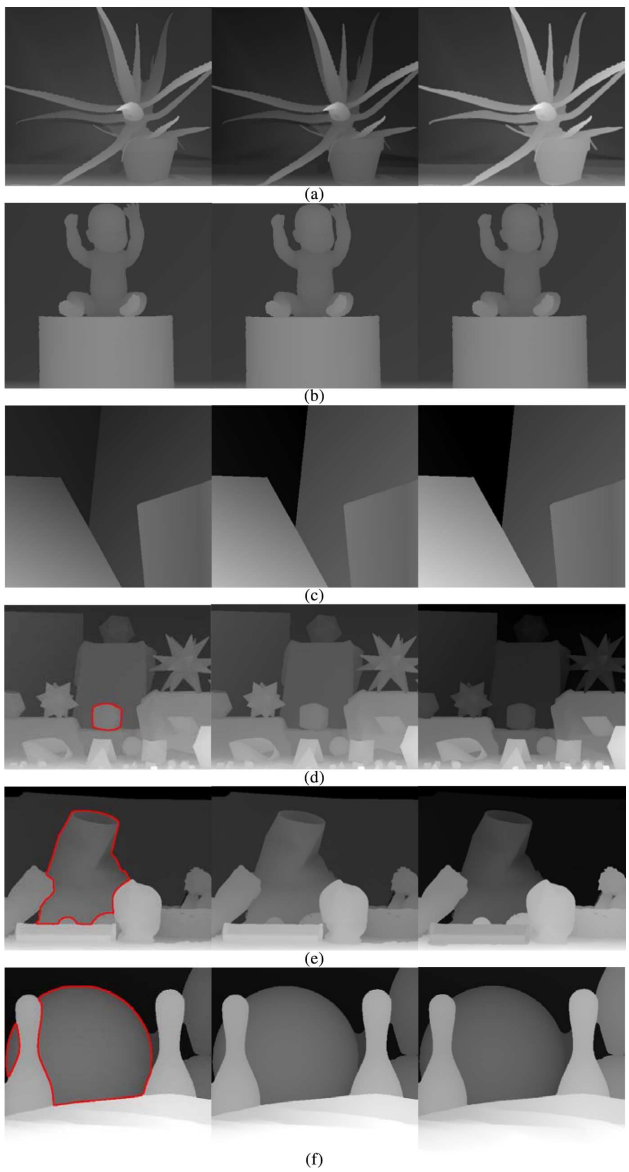


Fig. 8. Results of depth sensation enhancement. The depth maps from top to bottom are from the images of (a) Aloe, (b) Baby1, (c) Venus, (d) Moebius, (e) Middl1, and (f) Bowling1. From left to right are original depth maps (left), resultant depth maps using the conventional method [20] (middle), and resultant depth maps using the proposed method (right).

red border lines, are the salient objects in each corresponding scene. The depth differences between the salient objects and their neighboring objects are obviously stretched.

In the next step, the enhanced depth map is adopted to render virtual view images. In the experiment, two intermediate views to be rendered are taken as example. In other words, *N* in (1) is set as 2.

Since an increase in the depth difference between segments may produce additional holes after image rendering, the holes should also be considered for the visual quality of the stereoscopic images. Thus, image inpainting is used for filling holes [37] after virtual view rendering. Fig. 9(a) and Fig. 10(a) show the results of rendering without hole filling for Venus and Moebius, respectively, and Fig. 9(b) and Fig. 10(b) show the hole filling results. It can be observed that, the hole filling processing is important to ensure the quality of the generated stereoscopic images.

C. Subjective Evaluation of Depth Sensation and Image Quality on 3-D Display

In this section, we compare the depth sensation and the visual quality of the stereoscopic images generated by the original depth map, the conventional algorithm [20], and the proposed algorithm. For fair comparison, the energy functions in all comparing algorithms were solved by genetic algorithm with the same crossover, mutation operations, and the same parameter settings. Subjective evaluations were conducted using the 3-D WINDOWS-19A0 display produced by Tianjin 3-D Imaging Technique Co., Ltd. The environmental illuminance was 200 lux, the peak brightness was 250 cd/m^2 and the black level brightness was 0.05 cd/m^2. We evaluated the subjective quality of the generated stereoscopic images using stimulus comparison [38]. Twenty non-expert subjects with normal or corrected-to-normal visual acuity participated in the experiment. All of them had no experience with depth sensation enhancement. The subjects were asked to assess the visual quality and depth sensation while watching the content on the 3-D display. Based on the stimulus comparison method [20], [38], the subjects marked according to the comparison scale in Table I. Firstly, the stereoscopic images were composited by the left view and the 1/3-viewpoint virtual image. It can be observed that the proposed algorithm improves the depth sensation for the stereoscopic image contents and outperforms the conventional algorithm [20], as shown in Fig. 11. The error bars indicate 95% confidence interval. Since the depth values are properly changed, the stereoscopic images obtained using the proposed algorithm provide better depth sensation than those obtained using the conventional algorithm. The paired t-test [30] was performed to evaluate the statistical significance of the subjective evaluation re-

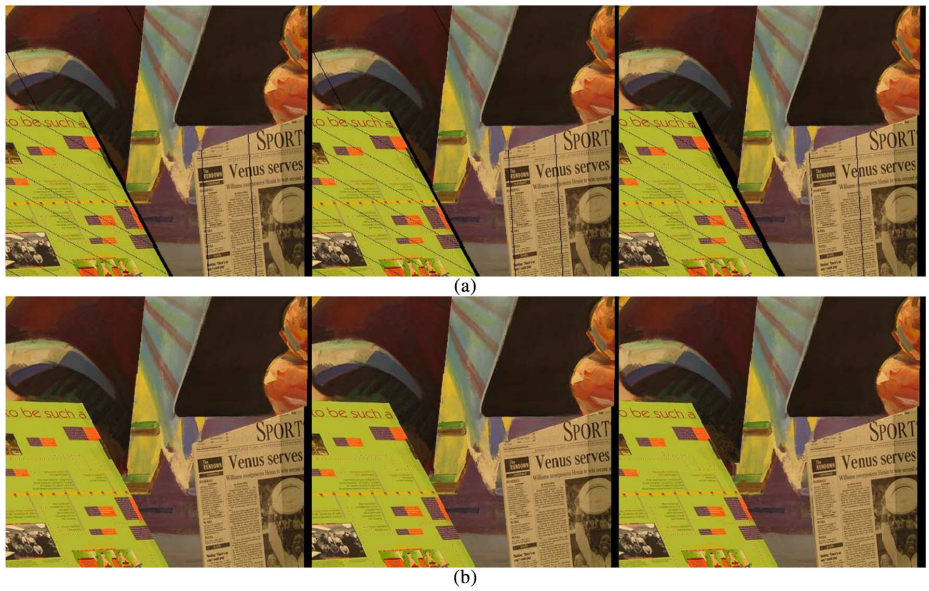


Fig. 9. Image rendering results from Venus. (a) 1/3-viewpoint images using the original depth map (left), the depth map obtained using the conventional method (middle), and the depth map obtained using the proposed method (right), respectively. (b) 1/3-viewpoint images after hole fillings for (a).

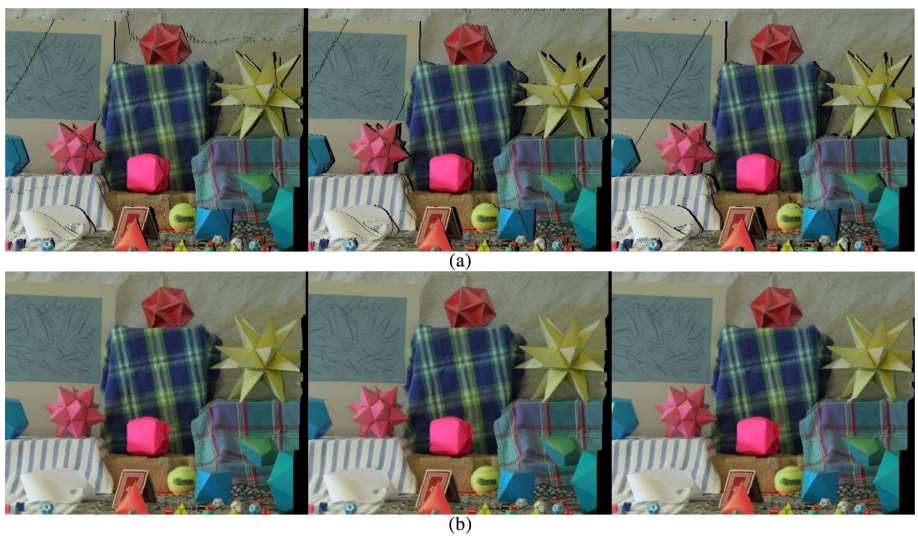


Fig. 10. Image rendering results from Moebius. (a) 1/3-viewpoint images using the original depth map (left), the depth map obtained using the conventional method (middle), and the depth map obtained using the proposed method (right), respectively. (b) 1/3-viewpoint images after hole fillings in (a).

sults. The computed t-value for Fig. 11(a) is 0.0019. When a universal threshold of 0.05 is used, the comparison result is statistically significant.

In order to show that the proposed method can be commonly applied for other virtual viewpoints, the same subjective evaluation experiment for stereoscopic display was conducted using the 1/3- and 2/3- virtual viewpoints. As shown in the experimental results from Fig. 12(a), the proposed algorithm outperforms the conventional algorithm, and improves the depth sensation of the stereoscopic image. The computed t-value for Fig. 12(a) is 0.0059, which means the comparison result is statistically significant.

It is worthwhile to mention that, the evaluation results of Baby1 in Figs. 11 and 12 are almost the same between different algorithms and the depth sensation scores are close to zero, since the depth differences in Baby1 are large enough. In other words, the depth differences between adjacent objects in the original depth map of Baby1, are larger than MVR_{JNDD} and do not need to stretch. This phenomenon is also consistent with the

TABLE I
COMPARISON SCALE FOR THE SUBJECTIVE QUALITY EVALUATION

-3	Much worse
-2	Worse
-1	Slightly worse
0	The same
+1	Slightly better
+2	Better
+3	Much better

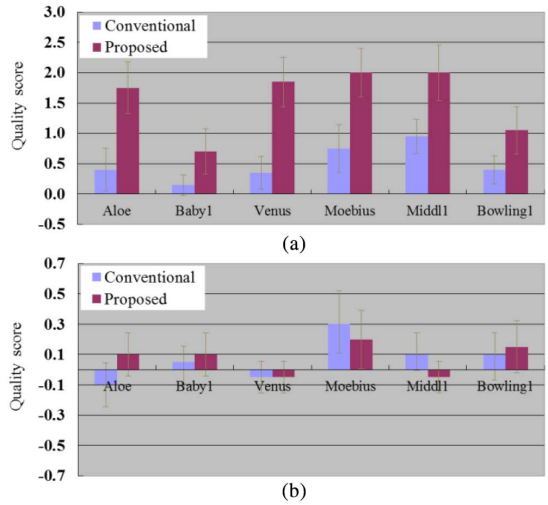


Fig. 11. Subjective quality evaluation results using the left viewpoint and the 1/3-virtual viewpoint. (a) Depth sensation. (b) Visual quality.

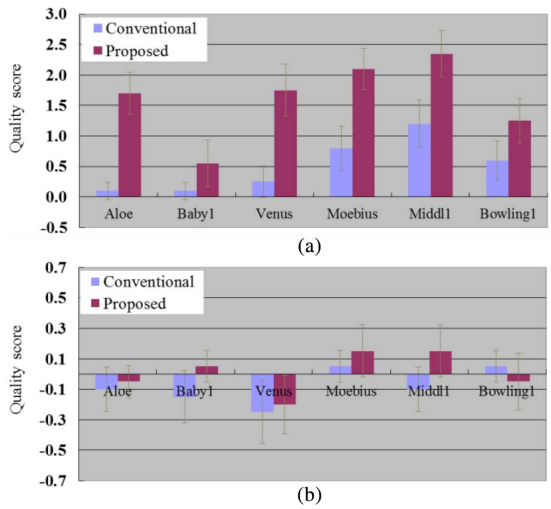


Fig. 12. Subjective quality evaluation results using the 1/3-virtual viewpoint and the 2/3-virtual viewpoint. (a) Depth sensation. (b) Visual quality.

objective experimental data. It is confirmed that depth sensation enhancement method would not change the depth map if it is not required to be enhanced.

For simple scenes, since the main difference is the threshold used in the process of depth sensation optimization, the complexity of the proposed method does not increase in this case. For complex scenes, the saliency analysis is adopted to give preferred guarantee to the depth differences between the salient object and its neighbors, thus the proposed method requires more time than the conventional method. However, the running time of the saliency extraction [29] is much shorter than the process of depth sensation optimization. Therefore, there is little increase in terms of complexity for the proposed method.

D. Evaluation on Multiview Autostereoscopic Display

In order to sufficiently evaluate the performance of our proposed method, a multiview autostereoscopic display called 3DFreeEyE42HD based on parallax barrier was used in the experiment. The same original color images (as shown in Fig. 5) were adopted. Each reference view and its corresponding enhanced depth map are used to render seven views (linearly change c in (2) from 1/7 to 7/7). Then, eight views (the reference view and seven virtual views) of each scene are obtained as input multiview images. For comparison, the original depth maps and the resultant depth maps using the conventional method [20] are also used to generate the resultant images for autostereoscopic display. Fig. 13 shows the subjective quality evaluation results with the same evaluation standard described in Section IV-C. In this experiment, the resultant stereoscopic images are displayed on the multiview autostereoscopic display. As shown in the experimental results in Fig. 13, the proposed algorithm improves the depth sensation of the multiview images for autostereoscopic display. The computed t-value for Fig. 13(a) is 0.0024, which indicates that the comparison result is statistically significant.

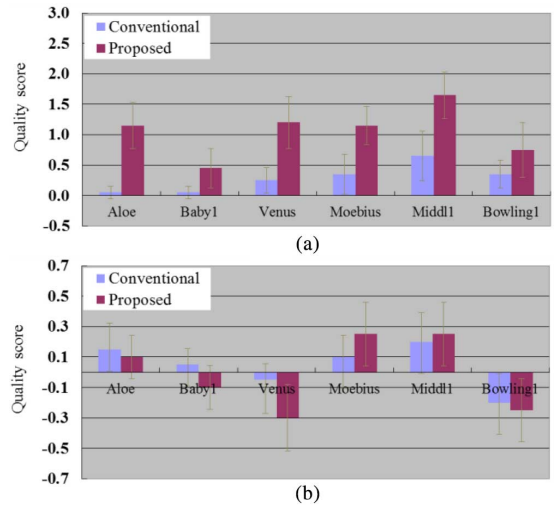


Fig. 13. Subjective quality evaluation results for multiview autostereoscopic display. (a) Depth sensation. (b) Visual quality.

E. Evaluation of Depth Sensation Enhancement With Video

To evaluate the performance of our depth enhancement method for video, we test two sequences, Alt Moabit [39] and Lovebird1 [40]. The corresponding original depth videos are generated by Depth Estimation Reference Software (DERS) [41]. The number of depth frames was set to 50. The original views of the two test sequences and their corresponding depth sequences are illustrated in Fig. 14. Firstly, the manual segmentation method is adopted. The segmentation results are shown in Fig. 15(a). The enhancement results for multiview autostereoscopic display with temporal term constraint and without temporal term constraint are shown in Fig. 15(c) and (d), respectively. Then, the fully-automatic segmentation method in [28] is used, and the experimental results are shown in Figs. 16 and 17. From the figures, it can be observed that the depths of the silver car in Alt Moabit (marked in Figs. 15 and 16) are more consistent between different frames with the proposed temporal term. Similarly, the depths of the tree in Lovebird1 (marked in Fig. 17) processed with the temporal term have better temporal consistency.

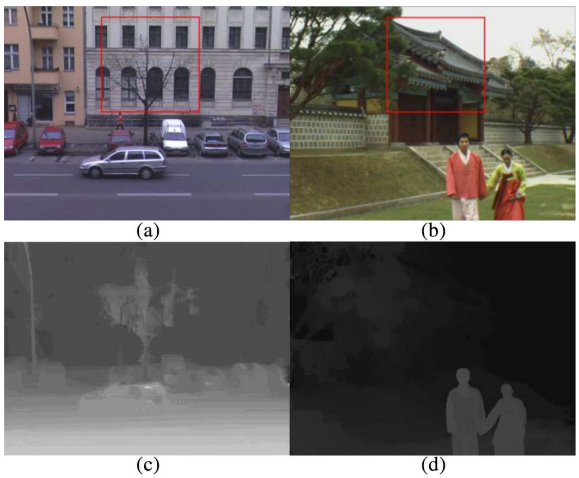


Fig. 14. Test sequences (a)–(b) and their corresponding depth maps (c)–(d). (a) Alt Moabit. (b) Lovebird1.

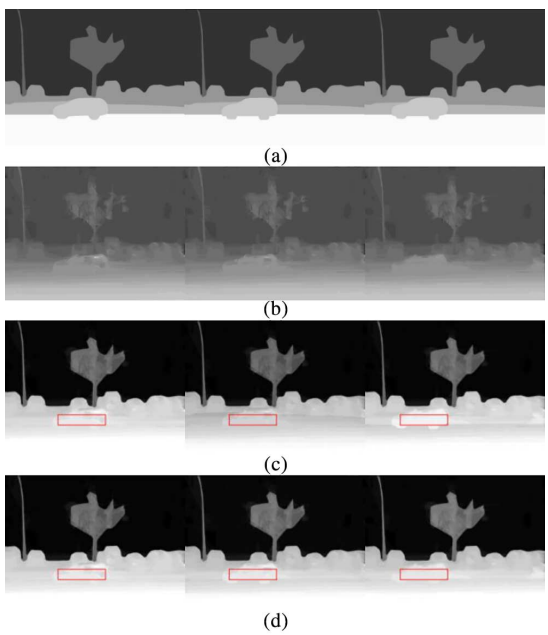


Fig. 15. Enhancement results of the 1st–3rd frames of Alt Moabit by semi-automatic method. (a) Manual segmentation results. (b) Original depth maps. (c) Enhancement results without temporal term. (d) Enhancement results with temporal term.

To objectively evaluate the temporal consistency, we analyzed the degree of depth fluctuation by calculating the average depth values in static areas which are marked in Fig. 14. The average depth values of marked regions are plotted in Fig. 18. The dotted line and solid line represent the results without temporal term and with temporal term, respectively. It can be observed from the figure that the depth fluctuation is strongly reduced with the temporal consistency preservation term. The standard deviations are 17.19 without temporal term constraint and 7.16 with temporal term constraint for the marked region of Alt Moabit, and 5.80 and 1.21 for Lovebird1, respectively. The proposed method can improve the temporal consistency of depth sensation enhancement for depth videos.

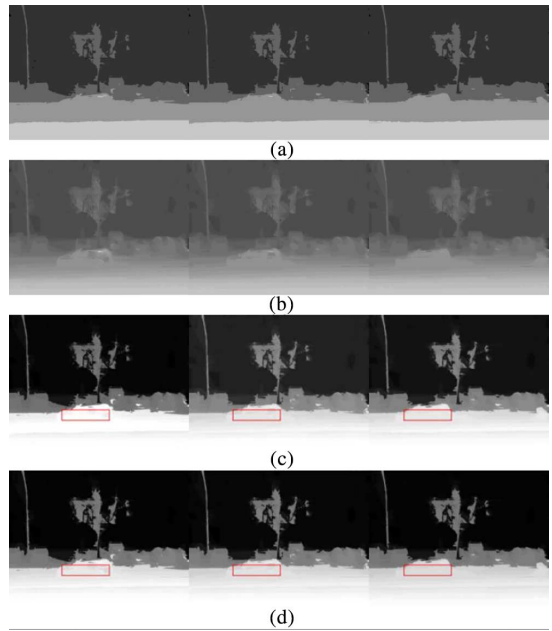


Fig. 16. Enhancement results of the 1st–3rd frames of Alt Moabit by fully-automatic method. (a) Fully-automatic segmentation results. (b) Original depth maps. (c) Enhancement results without temporal term. (d) Enhancement results with temporal term.

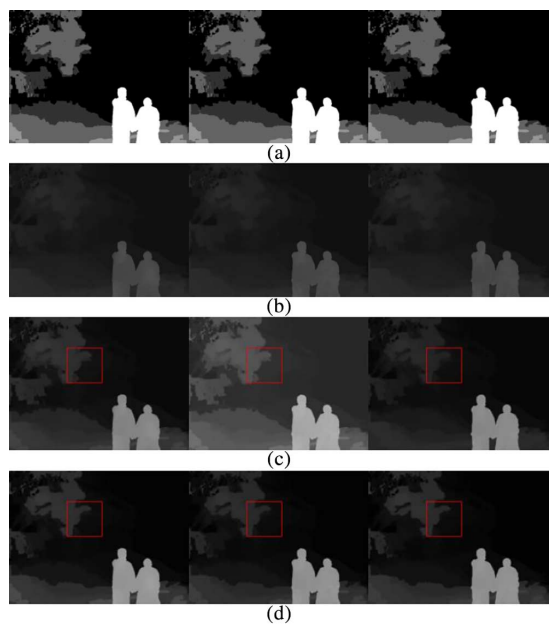


Fig. 17. Enhancement results of the 36th–38th frame of Lovebird1 by fully-automatic method. (a) Fully-automatic segmentation results. (b) Original depth maps. (c) Enhancement results without temporal term. (d) Enhancement results with temporal term.

F. Discussion

It should be noted that the proposed algorithm sometimes obtains slightly lower visual quality scores compared to those obtained using the original depth map. This is because that the hole regions usually increase with an increase of the depth difference, and precise hole filling is not always guaranteed. Generally, there are two basic types of 3-D video formats for DIBR, namely the single-view-plus-depth (SVD) format and the multi-view-

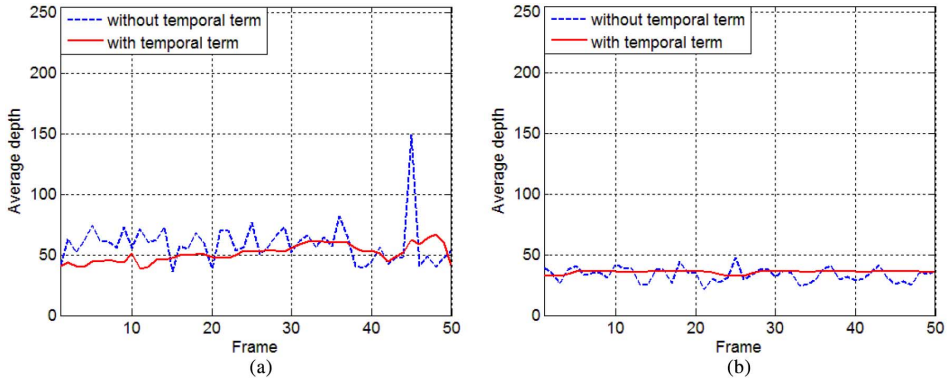


Fig. 18. Average depth variation of the marked region. (a) Alt Moabit. (b) Lovebird1.

plus-depth (MVD) format [42]. The proposed depth sensation enhancement method adopts SVD as input in this study. MVD, including stereo color-depth and multiview color-depth, may be available as input for multiview 3-D display. If the possible degradation of the visual quality of the 3-D contents is unacceptable in some applications, we can try to use image blending technologies based on MVD. In this situation, both the left-view and right-view depth maps are required to be enhanced, and the processed right-view depth map has to be matched with the enhanced left-view depth map.

It is also worthwhile to mention that depth segmentation is the basis of depth sensation enhancement. A better fully-automatic depth segmentation algorithm is desired for the promotion and application of depth sensation enhancement.

V. Conclusion

In this paper, we have presented a depth sensation enhancement method for multiple virtual view rendering. Different from previous studies, the proposed method considers the influence of virtual multiview rendering and the complexity of the scene. First, we propose to enhance the depth sensation of objects by considering the number of rendering intermediate views. Besides, the saliency measure is adopted to guarantee the depth sensation enhancement of salient objects in images. Then, the depth sensation enhancement is conducted based on a proposed energy function considering the number of rendering views and the selective visual attention. Finally, considering the temporal consistency between adjacent frames, a novel energy function with temporal consistency preservation term is proposed to extend the depth sensation enhancement to video applications. Experimental results demonstrate that the depth perception is successfully improved as expected. The proposed method makes it possible to apply depth sensation enhancement for 3-D display based on DIBR.

Acknowledgment

The authors would like to thank Prof. M.-T. Sun and Dr. X. Ye for their valuable suggestions and kind help. The authors also would like to thank the anonymous reviewers for their valuable comments that significantly helped us in improving the presentation of the paper.

References

[1] P. Ndjiki-Nya *et al.*, "Depth image-based rendering with advanced texture synthesis for 3-D video," *IEEE Trans. Multimedia*, vol. 13, no. 3, pp. 453–465, Jun. 2011.

[2] N. S. Holliman, N. A. Dodgson, G. E. Favalora, and L. Pockett, "Three-dimensional displays: A review and applications analysis," *IEEE Trans. Broadcast.*, vol. 57, no. 2, pt. 2, pp. 362–371, Jun. 2011.

[3] B. Macchiavello, C. Dorea, E. M. Hung, G. Cheung, and W.-T. Tan, "Loss-resilient coding of texture and depth for free-viewpoint video conferencing," *IEEE Trans. Multimedia*, vol. 16, no. 3, pp. 711–725, Apr. 2014.

[4] M. M. Hannuksela *et al.*, "Multiview-video-plus-depth coding based on the advanced video coding standard," *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3449–3458, Sep. 2013.

[5] P. Gao and W. Xiang, "Rate-distortion optimized mode switching for error-resilient multi-view video plus depth based 3-D video coding," *IEEE Trans. Multimedia*, vol. 16, no. 7, pp. 1797–1808, Nov. 2014.

[6] A. Kubota *et al.*, "Multiview imaging and 3DTV," *IEEE Signal Process. Mag.*, vol. 24, no. 6, pp. 10–21, Nov. 2007.

[7] S. Li, J. Lei, C. Zhu, L. Yu, and C. Hou, "Pixel-based inter prediction in coded texture assisted depth coding," *IEEE Signal Process. Lett.*, vol. 21, no. 1, pp. 74–78, Jan. 2014.

[8] W. Sun, O. Au, L. Xu, Y. Li, and W. Hu, "Seamless view synthesis through texture optimization," *IEEE Trans. Image Process.*, vol. 23, no. 1, pp. 342–355, Jan. 2014.

[9] C. Fehn, "Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV," in *Proc. Int. Soc. Opt. Eng.*, 2004, vol. 5291, pp. 93–104.

[10] F. Shao, G. Jiang, W. Lin, M. Yu, and Q. Dai, "Joint bit allocation and rate control for coding multi-view video plus depth based 3-D video," *IEEE Trans. Multimedia*, vol. 15, no. 8, pp. 1843–1854, Dec. 2013.

[11] S.-W. Jung and S.-J. Ko, "Depth enhancement considering just noticeable difference in depth," *IEICE Trans. Fundam. Electron., Commun. Comput. Sci.*, vol. 95, no. 3, pp. 673–675, 2012.

[12] N.-Y. Jo, H.-G. Lim, S.-K. Lee, Y.-S. Kim, and J.-H. Park, "Depth enhancement of multi-layer light field display using polarization dependent internal reflection," *Opt. Express*, vol. 21, no. 23, pp. 28758–28770, 2013.

[13] R. Näsänen, T. Colomb, Y. Emery, and T. J. Naughton, "Enhancement of three-dimensional perception of numerical hologram reconstructions of real-world objects by motion and stereo," *Opt. Express*, vol. 19, no. 17, pp. 16075–16086, 2011.

[14] S. Yang *et al.*, "Discernible difference and change in object depth afforded by stereoscopic three-dimensional content," in *Proc. SPIE, Stereoscopic Displays Appl. XXIV*, 2013, pp. 86481C-1–86481C-11.

[15] H. Sohn, Y. J. Jung, and Y. M. Ro, "Local disparity remapping to enhance depth quality of stereoscopic 3-D images using stereoacuity function," in *Proc. SPIE, Stereoscopic Displays Appl. XXV*, 2014, pp. 90110L-1–90110L-6.

[16] G. Nur, S. Dogan, H. K. Arachchi, and A. M. Kondoz, "Impact of depth map spatial resolution on 3-D video quality and depth perception," in *Proc. 3DTV Conf. The True Vision - Capture, Transmiss. Display 3D Video*, 2010, pp. 1–4.

[17] S. Heinzle *et al.*, "Computational stereo camera system with programmable control loop," *ACM Trans. Graphics*, vol. 30, no. 4, pp. 1–10, 2011.
[18] D. V. S. De Silva, W. A. C. Fernando, S. T. Worrall, S. L. P. Yasakethu, and A. M. Kondoz, "Just noticeable difference in depth model for stereoscopic 3-D displays," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2010, pp. 1219–1224.
[19] D. V. S. De Silva, E. Ekmekcioglu, W. A. C. Fernando, and S. T. Worrall, "Display dependent preprocessing of depth maps based on just noticeable depth difference modeling," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 2, pp. 335–351, Apr. 2011.
[20] S.-W. Jung and S.-J. Ko, "Depth sensation enhancement using the just noticeable depth difference," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3624–3637, Aug. 2012.
[21] S.-W. Jung, "A modified model of the just noticeable depth difference and its application to depth sensation enhancement," *IEEE Trans. Image Process.*, vol. 22, no. 10, pp. 3892–3903, Oct. 2013.
[22] M. Z. Brown, D. Burschka, and G. D. Hager, "Advances in computational stereo," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 8, pp. 993–1008, Aug. 2003.
[23] K. Muller, P. Merkle, and T. Wiegand, "3-D video representation using depth maps," in *Proc. IEEE*, Apr. 2011, vol. 99, no. 4, pp. 643–656.
[24] S. M. Seitz and C. R. Dyer, "Physically-valid view synthesis by image interpolation," in *Proc. IEEE Workshop Representation Visual Scenes*, Jun. 1995, pp. 18–25.
[25] S. Wang and J. M. Siskind, "Image segmentation with ratio cut," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 6, pp. 675–690, Jun. 2003.
[26] L. T. Ha, S.-W. Jung, K.-S. Choi, and S.-J. Ko, "Image segmentation based on a modified graph-cut algorithm," *IET Electron. Lett.*, vol. 46, no. 16, pp. 1121–1123, 2010.
[27] B. Dellen, G. Alenyà, S. Foix, and C. Torras, "Segmenting color images into surface patches by exploiting sparse depth data," in *Proc. IEEE Workshop Appl. Comput. Vis.*, Jan. 2011, pp. 591–598.
[28] D. V. S. De Silva, W. A. C. Fernando, H. Kodikaraarachchi, S. T. Worrall, and A. M. Kondoz, "Adaptive sharpening of depth maps for 3D-TV," *Electron. Lett.*, vol. 46, no. 23, pp. 1546–1548, 2010.
[29] J. Lei, H. Zhang, L. You, C. Hou, and L. Wang, "Evaluation and modeling of depth feature incorporated visual attention for salient object segmentation," *Neurocomput.*, vol. 120, pp. 24–33, 2013.
[30] Y. Fang, J. Wang, M. Narwaria, P. L. Callet, and W. Lin, "Saliency detection for stereoscopic images," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2625–2636, Jun. 2014.
[31] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu, "Global contrast based salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Mar. 2011, pp. 409–416.
[32] Y. Fang, W. Lin, Z. Chen, C.-M. Tsai, and C.-W. Lin, "A video saliency detection model in compressed domain," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 1, pp. 27–38, Jan. 2014.
[33] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. J. Comput. Vis.*, vol. 47, no. 1–3, pp. 7–42, 2002.
[34] D. Scharstein and C. Pal., "Learning conditional random fields for stereo," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.
[35] H. Hirschmüller and D. Scharstein, "Evaluation of cost functions for stereo matching," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.
[36] X. Xu *et al.*, "A foreground biased depth map refinement method for DIBR view synthesis," in *Proc. IEEE Int. Conf. Speech Signal Process. Acoust.*, Mar. 2012, pp. 805–808.
[37] J. Zhou and A. Robles-Kelly, "Image inpainting based on local optimization," in *Proc. IEEE Int. Conf. Pattern Recognit.*, Aug. 2010, vol. 2010, no. 20, pp. 4440–4443.
[38] *Methodology for the Subjective Assessment of the Quality of Television Pictures*, ITU-R Recommendation BT.500-11, Int. Telecommun. Union, Geneva, Switzerland, 2002.
[39] *HHI Test Material for 3-D Video*, Doc. M15413, ISO/IEC JTC1/SC29/WG11, Archamps, France, Apr. 2008.
[40] *Contribution for 3-D Video Test Material of Outdoor Scene*, MPEG2008/M15371, ISO/IEC JTC1/SC29/WG11 , Archamps, France, Apr. 2008.
[41] *Reference Softwares for Depth Estimation and View Synthesis*, Doc. M15377, ISO/IEC JTC1/SC29/WG11, Archamps, France, Apr. 2008.
[42] C. Yao *et al.*, "Depth map driven hole filling algorithm exploiting temporal correlation information," *IEEE Trans. Broadcast.*, vol. 60, no. 2, pp. 394–404, Jun. 2014.

Jianjun Lei (A'11–M'12) received the Ph.D. degree in signal and information processing from Beijing University of Posts and Telecommunications, Beijing, China, in 2007.

He was a visiting Researcher with the Department of Electrical Engineering, University of Washington, Seattle, WA, USA, from August 2012 to August 2013. He is currently a Full Professor with the School of Electronic Information Engineering, Tianjin University, Tianjin, China. His research interests include 3-D video processing, 3-D display, and computer vision.

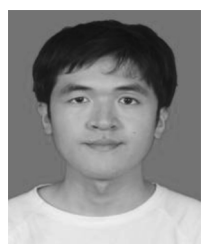
Cuicui Zhang received the B.S. degree in communication engineering from Tianjin University, Tianjin, China, in 2013, and she is currently working towards the M.S. degree at the School of Electronic Information Engineering, Tianjin University, Tianjin, China.

Her research interests include 3-D image processing and 3-D display.

Yuming Fang (M'13) received the B.E. degree from Sichuan University, Chengdu, China, the M.S. degree from Beijing University of Technology, Beijing, China, and the Ph.D. degree in computer engineering from Nanyang Technological University, Singapore.

He was a visiting Postdoc Research Fellow with the IRCCyN Laboratory, PolyTech Nantes, Nantes, France, the University of Nantes, Nantes, France, the University of Waterloo, Waterloo, ON, Canada, and Nanyang Technological University, Singapore. He is currently a Faculty Member with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang, China. His research interests include visual attention modeling, visual quality assessment, image retargeting, computer vision, and 3-D image/video processing.

Dr. Fang was a Secretary for the 9th Joint Conference on Harmonious Human Machine Environment. He was also a Special Session Organizer for VCIP 2013 and QoMEX 2014.



Zhouye Gu (M'14) received the B.Sc. degree in electronic information science and technology from Nanjing University, Nanjing, China, in 2007, and the Ph.D. degree in computer engineering from Nanyang Technological University, Singapore, in 2014.

He was a Visiting Student with the University of Washington, Seattle, WA, USA, in 2012, and a Research Fellow with Santa Clara University, Santa Clara, CA, USA, from 2012 to 2014. He is currently a Staff Advanced Research Engineer with ARRIS Group Inc., CA, USA. His research interests include video coding and processing, machine learning, and data mining.

Dr. Gu was the Technical Program Co-Chair of U-Media 2014 and the Area Chair of ICME 2013.

Nam Ling (S'85–M'88–SM'99–F'08) received the B.Eng. degree from the National University of Singapore, Singapore, in 1981, and the M.S. and Ph.D. degrees from the University of Louisiana, Lafayette, LA, USA, in 1985 and 1989, respectively.

From 2002 to 2010, he was an Associate Dean with the School of Engineering, Santa Clara University, Santa Clara, CA, USA. He is currently the Sanfilippo Family Chair Professor and the Chair for the Department of Computer Engineering, Santa Clara University. He is currently also a Consulting Professor with the National University of Singapore, a Guest Professor for Shanghai Jiao Tong University, Shanghai, China, a Cuiying Chair Professor for Lanzhou University, Gansu, China, an Outstanding Overseas Scholar for the Shanghai University of Electric Power, Shanghai, China, and a Distinguished Professor for Xi'an University of Posts and Telecommunications, Shaanxi, China. He has authored or coauthored over 170 publications and standard contributions, including two books in the fields of video coding and systolic arrays. He has filed/granted over 10 U.S. patents.

Dr. Ling is an IET Fellow. He was named as an IEEE Distinguished Lecturer twice and is currently an APSIPA Distinguished Lecturer. He received the IEEE ICCE Best Paper Award (First Place). He was a recipient of six awards from Santa Clara University, four at the University level (Outstanding Achievement, Recent Achievement in Scholarship, President's Recognition, and Sustained Excellence in Scholarship) and two at the School/College level (Researcher of the Year and Teaching Excellence). He was a Keynote Speaker for IEEE APCCAS, VCVP (twice), JCPC, IEEE ICAST, IEEE ICIEA, IET FC Umedia, and IEEE Umedia, as well as a Distinguished Speaker for IEEE ICIEA. He has served as a General Chair/Co-Chair for IEEE Hot Chips, VCVP (twice), IEEE ICME, IEEE Umedia (twice), and IEEE SiPS. He has also served as a Technical Program Co-Chair for IEEE ISCAS, APSIPA ASC, IEEE APCCAS, IEEE SiPS (twice), DCV, and IEEE VCIP. He was a Technical Committee Chair for IEEE CASCOM TC and IEEE TCMM, and has served as a Guest Editor or Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: REGULAR PAPERS, the IEEE JOURNAL OF SELECTED TOPICS N SIGNAL PROCESSING, JSPS, and Springer *MSSP*.

Chunping Hou received the M.Eng. and Ph.D. degrees in electronic engineering from Tianjin University, Tianjin, China, in 1986 and 1998, respectively.

Since 1986, she has been with the School of Electronic and Information Engineering, Tianjin University, Tianjin, China, where she is currently a Full Professor and the Director of the Broadband Wireless Communications and 3-D Imaging Institute. Her current research interests include 3-D image processing, 3-D display, wireless communication, and the design and applications of communication systems.